



## Artificial societies. 2013-2020

ISSN 2077-5180

URL - <http://artsoc.jes.su>

All right reserved

Issue 4 Volume 14. 2019

# Ethical guidelines for intellectual technologies in education

**A. Gorman**

*Moscow state university  
Russian Federation, Moscow*

## Abstract

If we put the well-being of each individual student above all, it is obvious that the use of artificial intelligence (AI) can seriously improve the situation with the quality and equality of access to education in the long term. However, this is possible if and only if in the process of implementation and application of such systems, the rights and freedoms of every person will not be violated, which requires appropriate upbringing and education of citizens, consensus on ethical standards, their implementation at the regulatory intergovernmental, technological and economic levels, as well as directed preventive technological solutions.

**Keywords list (en):** artificial intelligence, technology ethics, education, educational technology

**Date of publication:** 19.12.2019

## Citation link:

Gorman A. Ethical guidelines for intellectual technologies in education // Artificial societies. 2019. V. 14. Issue 4 [Electronic resource]. Access for registered users. URL: <https://artsoc.jes.su/s207751800007918-8-1/> (circulation date: 28.09.2020). DOI: 10.18254/S207751800007918-8

1

## ИИ способен решить текущие проблемы образования

Искусственный интеллект при грамотном применении способен решить основные проблемы образования: неравный доступ к образованию, психологический дискомфорт участников образовательного процесса, неспособность текущей образовательной системы

подготовить человека к успеху в борьбе с вызовами будущего.

## 2 (1) Неравный доступ к образованию

3 По данным отчета Economic Policy Institute, множественные исследования показали, что социальный класс учащегося является самым важным фактором успеха образовательного процесса [6], в силу того, что тьюторство, доступное немногим, является наиболее эффективным методом преподавания, так как учитывает индивидуальные особенности ученика. Уже сейчас, согласно Гоэль и Жоине [10], основанный на технологии искусственного интеллекта помощник, построенный на основе платформы Watson от IBM, используется в массовом открытом онлайн-курсе Knowledge Based Artificial Intelligence института технологии Georgia Institute of Technology для ответа на вопросы слушателей курса, что позволяет уже предоставить не только широкий доступ к образовательным материалам, но и создавать условия для персонализированного обучения.

4 Если избавление от когнитивных искажений у преподавателей в процессе преподавания и оценки учеников и студентов потребует смены типа мышления, а также длительных социальных изменений, то технологически достижение, например, гендерного равенства, произойдет значительно быстрее при применении систем ИИ, обучающихся без участия данных человека. При этом, эффективность программ, не использующих данные человека, значительно выше: так, AlphaGo Zero алгоритм от компании Deepmind обыграл в Го все предыдущие свои версии, обученные на данных из прошлого.

5 Применение систем искусственного интеллекта в образовании может позволить решить актуальные проблемы специальной педагогики, в том числе согласно принципам инклюзивного образования, в рамках которого требующие отдельного внимания, в силу особенностей развития обучающиеся могут получить более широкие возможности для последующей адаптации в социуме. Так, например, Фрауенбергер и др. [8], отмечают в своем исследовании мультидисциплинарной методики Echoes, что системы искусственного интеллекта помогают развивать, в том числе и социальные навыки детей с особенностями развития, например детей, страдающих аутизмом.

## 6 (2) Психологический дискомфорт участников образовательного процесса

7 Многие из решений данной проблемы перекликаются с решениями, направленными на предоставление равного доступа к системе образования, так как они основаны на персонализированном обучении, учитывающем, в том числе, когнитивную загруженность ученика, а также его психоэмоциональные особенности. Применение ИИ может учитывать когнитивную нагрузку учащихся, предотвращая сам источник стресса. Также ИИ может использоваться для раннего обнаружения проблем буллинга в рабочих группах учеников и студентов, для создания образовательной среды, психологически комфортной для всех участников образовательного процесса.

8 Преподаватель в классической образовательной среде не имеет достаточных ресурсов и часто достаточного набора навыков для учета психоэмоционального состояния каждого учащегося. Согласно Хуангу и Хвангу [12], учитывая достижения аффективного программирования, исследователи также могут объективно определять и измерять эмоциональный статус учащегося в реальном времени на протяжении всего образовательного процесса, чтобы лучше понять связь между эмоциями, мотивацией и образовательными результатами.

9 В медицине, в области ухода за тяжелобольными пациентами, уже применяются роботы различного вида для облегчения психологического стресса пациентов. Например, согласно исследованию Кало и др [4], применение робота-тюленя компании Takanoti Shibata

показало существенный терапевтический эффект. Нет никаких оснований полагать, что использование аффективного программирования в области образования для снятия психоэмоционального напряжения будет обладать меньшим эффектом.

10 (3) Неспособность текущей образовательной системы подготовить человека к успеху в борьбе с вызовами будущего складывается, во-первых, из применения недоказательных педагогических методов, а также ориентации текущей системы образования исключительно на академические успехи.

11 Применение искусственного интеллекта может дать новый виток в развитии доказательной педагогики – то есть педагогики, согласно Харгривсу [11], основанной на методиках доказанной эффективности, с использованием последних достижений в науке и технологиях. Так, например, без систем искусственного интеллекта проблема объективной оценки успеваемости на данный момент решена не была. Помимо оценки непосредственно успеваемости, системы искусственного интеллекта могут оценивать отдельные драйверы успеваемости. Так, согласно Лесгольд и др. [16], методики с использованием искусственного интеллекта “могут определять эффективность эпизодов обмена знанием между студентами”, что помогает “определить ситуации, в которых фасилитация может повысить эффективность групповых взаимодействий”. Такие “ситуации” будут иметь доказанную эффективность, подтвержденную научными методами, и поэтому смогут применяться в более общей педагогической практике, что даст возможное развитие области наук об образовании.

12 По результатам мета-исследования литературы Лаар и др [13], были выделены ключевые навыки для успеха в будущем: технологические навыки, навыки управления информацией, коммуникативные, навыки совместной деятельности, творчества, критического мышления и решения проблем. Кроме этого, были добавлены следующие навыки: осведомленность об этике, культуре, гибкость, навыки самообучения и непрерывного обучения. Очевидно, что многие из этих навыков требуют пересмотра текущей системы образования и добавления технологических аспектов.

13 Таким образом, во многом, применение ИИ решило бы многие этические вопросы, которые стоят перед текущей системой образования. Это и снижение количества “образовательных опытов” в медицине, и развитие меритократии, и общее улучшение качества образования, что соответственно ведет к росту качества жизни большего количества людей. Однако применение ИИ ставит и новые вызовы, основными из которых являются: защита персональных данных, трансформация рынка труда, изменения в восприятии роли обучения и учителя, психологические риски взаимодействия с машинами и формирование систем, не создающих новых “хранителей”.

14

## **Барьеры и сложности на пути применения систем ИИ**

Применение алгоритмов и моделей искусственного интеллекта в образовании заставляет преподавателей и других участников образовательного процесса пересмотреть свои предпосылки, суть собственной профессиональной деятельности. Согласно Ричарду и Даниэлю Зюскинд [17], профессия преподавателя уже трансформируется от работы в виде “мудреца на сцене” к становлению “проводниками-помощниками”, то есть роль преподавателя может быть в будущем сведена к ассистированию обучающимся в выборе систем и методик, что оказывает существенное влияние на взаимоотношения участников процесса обучения, а также на всю сферу образования. Как и в других профессиях, общество можно столкнуться как с технологической безработицей, так и другими последствиями широкого применения машин с расширяющимися возможностями, что требует детального рассмотрения этих влияний с философской точки зрения.

15 Такие исследования, как успешное тестирование электростимуляции мозга для улучшения образовательных результатов у пилотов, проведенное в HRL лаборатории [5], заставляют участников образовательного процесса задуматься о природе знаний, их получении и будущей востребованности. Трансформируется не только роль учителя, который уже перестает выступать источником знаний, но меняется и восприятие знаний, их ценности, общее мировоззрение участников образовательного процесса.

16 Кроме этого, все больше указывает на то, что знания в классическом понимании уже не являются тем, что будет определять благополучие членов общества в будущем. Так, например, мастер игры в Го из Южной Кореи Ли Си Дол уже признал превосходство ИИ и оставил карьеру профессионального игрока [19]. Будет ли каждый отдельный человек ощущать когнитивное бессилие перед мощностью “мозга” ИИ или найдет новое место и применение своим способностям?

17 Актуальность проблемы технологизации образования была замечена и ранее, Лаура и Чапман [14] отметили, что “когда молодые и немолодые люди, вовлекаются в удаленные взаимоотношения, посредством компьютерного киберпространства, создается иллюзия, что эти отношения наполненные и глубокие, в то время как на самом деле они являются лишь одномерной частью многомерной формы человеческого взаимодействия”. Данный вывод можно экстраполировать и на взаимодействия обучающихся в образовательной среде, роль которых и влияние на процесс обучения не до конца изучена.

18 Технологический прогресс не стоит на месте, но всегда встречается с сопротивлением. Ссопротивление любым изменениям начинается с отрицания. Однако стратегия отрицания встретится с реальностью, когда ученики, например, будут использовать алгоритм GPT2 от OpenAI для написания эссе, тексты которого уже сейчас в большинстве случаев неотличимы от написанных человеком.

19 Помимо сопротивления на уровне участников образовательного процесса, существует и риск развития иных конфликтов в обществе. Резкий скачок в уровне образования следующего поколения может формировать межпоколенческие социальные конфликты и еще большую разобщенность общества. При неравном доступе к новым технологиям, разница уровня дохода между защищенными и незащищенными слоями населения вырастет еще сильнее, что приводит к увеличению социальной напряженности. И, конечно же, не останутся в стороне и религиозные организации, которые могут сопротивляться новой роли технологий в образовании и воспитании общества. При этом стоящие перед системой образования вызовы в сфере будущего применения ИИ перекликаются с вызовами других областей, особенно в области защиты данных и прав человека, но имеют и специфические трудности, требующие отдельного рассмотрения.

20

## **Что может определить этический вектор развития применения ИИ в образовании?**

21 **Гипотеза 1** – Этические нормы применения технологий будут определяться этическими нормами и уровнем образования в этой области создателей технологий.

22 Технологии более не являются способом упрощения деятельности в отдельных специальностях или средством автоматизации процессов. Искусственный интеллект, как алгоритм, является живым организмом, отражающим мнения, осведомленность и этические нормы его создателей. Поэтому этика технологий начинается с педагогических усилий по воспитанию этических норм. Таким образом, роль будущих педагогов напрямую зависит от

текущих усилий и направлений в образовательном процессе.

23 Одной из возможностей решения потенциальных вызовов является обучение разработчиков, программистов, ученых по данным, специалистов по робототехнике и смежных специальностей основам этики. Сложность данного подхода состоит в том, что с одной стороны, обучение должно оказаться достаточным для разработки соответствующих решений, с другой – полный охват дисциплин философии и, в частности, этики – может быть чрезмерным с практической точки зрения, и поэтому нереалистичным. В рамках фундаментального образования в России, несмотря на наличие предмета философии на разных учебных курсах, на выходе специалисты зачастую не обладают достаточной подготовкой к принятию сложных этических решений.

24 Фурей и Мартин [9] предложили модульный подход к обучению этике в области искусственного интеллекта, в рамках которого они включили модуль по этике в курс по искусственному интеллекту, используя проблемный метод при обсуждении проблем разработки алгоритмов для беспилотных автомобилей с помощью известной задачи с вагонетками. Исследователи отметили, что обсуждения помогли студентам оценивать потенциальные положительные и негативные стороны интуитивного решения этических проблем – утилитарного подхода. Кроме этого, исследователи отметили, что подобный эксперимент позволил познакомить студентов с ключевым навыком решения этических проблем – использованием “мысленного эксперимента”. Сложность этической проблемы, представленная в параллели со сложностью технического решения, помогает студентам учитывать этические аспекты при разработке.

25 В этом подходе, создание практико-ориентированных материалов для обучения разработчиков, предпринимателей и других людей, которые принимают решения о будущих алгоритмах, стоящих за машинами с искусственным интеллектом, требует совместных усилий педагогов и философов.

26 В MIT [15] придерживаются подхода внедрения и более раннего обучения этике искусственного интеллекта. Один из проектов уже разработал и выложил в открытый доступ расписание для учеников средней школы, призванное обучить студентов технологическим концепциям и их этическим последствиям с конечной целью снабдить студентов инструментами и навыками разработки ИИ с учетом этических норм.

27 **Гипотеза 2** – Этические нормы применения технологий будут определяться глобальным консенсусом

28 Другой возможностью решения потенциальных вызовов может быть создание специализированных центров и лабораторий, объединяющих специалистов различных областей для решения связанных с ИИ этических вопросов. Такого рода специализированные центры могут формировать общественные дискуссии при возникновении противоречий с существующими нормами. Возможно, впервые такой подход был применен в Великобритании после открытия искусственного оплодотворения, когда философы, среди прочих экспертов, подготовили отчет об этической составляющей нового тогда открытия.

29 Примером подобного сообщества может быть AI Ethics Lab [1], расположенного в Бостоне и Стамбуле. В сообщество входят исследователи и практики в области компьютерных наук, юристы, академики и философы. Компания использует метод мэппинга – структурного метода решения этических вопросов – при разработке дизайна технологических решений. Например, одним из последних результатов работы является разработка плана исследований по регулированию применения ИИ в области медицины, подготовленного с участием ученых из Гарварда и других исследовательских центров [3].

30 Другим примером неправительственной организации может быть основанный в Великобритании The Institute for Ethical AI & Machine Learning [18]. Этот институт фокусируется на 4 составляющих стратегии: принципы разработки, процесс, стандарт индустрии, регулятивные нормы. Или некоммерческая организация Algorithmwatch[2], продвигающая подход прозрачности в отношении используемых алгоритмов и публикующая открыто данные о том, как могут использоваться уже внедренные алгоритмы. Из последних исследований опубликованы отчеты о системах рейтинга граждан в Африке и Китае, а также последних изменений в политике социальной сети Facebook.

31 Согласно директиве ЕС по этике в ИИ [7], ставящая в центр законопроектов человека стратегия ЕС отличается от стратегии США, разработанной преимущественно через инициативы частного сектора и инструменты саморегуляции, и стратегии Китая, преимущественно с инициированными правительством мерами. Однако учитывая схожесть перечисленных в документе общих этических принципов и сложности межгосударственного контроля технологий, разработка универсальных правил могла бы более эффективной.

32 Более того, перечисленные в документе меры и нормы никак не касаются отдельных областей применения, в частности образования. Как показала история законов роботики Азимова, лишь применение частных случаев, мысленный эксперимент или ошибки на практике показывают всю сложность применения не известных ранее технологий в конкретных областях. Допустимо ли создание всеобщего рейтинга обучающихся? По каким критериям может быть составлен рейтинг и на что он может влиять? Возможно ли раскрытие данных о когнитивных возможностях работодателям при отборе кандидатов на продолжение профессионального обучения? Остается ли за человеком выбор траекторий развития?

33 В свете стоящих перед рынком труда вызовов, также имеются решения экономико-политического характера. Уже сейчас проводятся эксперименты по предоставлению безусловного дохода, а также возможностей людей монетизировать свои собственные данные.

34 **Гипотеза 3** – Этические нормы применения технологий будут определяться превентивными технологиями

35 Учитывая глобальный характер последующих вызовов, меры должны обладать менее централизованной структурой. Здесь сами технологии предоставляют возможности для децентрализации. Используя блокчейн-технологии, возможно добиться не просто оптимизации администрирования образования, но также защиты персональных данных на более высоком уровне. Таким образом, помимо создания законов и директив, мы можем начать с технологических решений, способных предотвратить появление некоторых этических проблем. По аналогии с медициной, до разработки лекарств и методик лечения гипотетического вируса, мы можем полностью остановить ее распространение, не дожидаясь случая первого заражения. Превентивными мерами могут быть создание децентрализованных блокчейн-систем с полным шифрованием, которые, например, могли бы защищать данные о когнитивных способностях учащихся, не дожидаясь того, что кто-то из работодателей или иных заинтересованных лиц может начать использовать данные для улучшения производительности бизнеса. Необходимо убрать материальную мотивацию взлома баз данных, то есть хакерские атаки должны быть дороже любой прибыли, возможной в результате взлома. По образовательным вопросам может быть также применен метод квадратичного голосования для решения общих вопросов, что позволило бы создать системы образования, действительно ориентированные на человека.

1. AI Ethics Lab // URL: <http://aiethicslab.com/>
2. Algorithm Watch // URL: <https://algorithmwatch.org/en/>
3. Arney D, Senges M, Gerke S. , A User-Focused Transdisciplinary Research Agenda for AI-Enabled Health Tech Governance // Cyber Harvard, January 2019 URL: [https://cyber.harvard.edu/sites/default/files/2019-02/2019-01\\_aihealth.pdf?subscribe=Download+the+Report](https://cyber.harvard.edu/sites/default/files/2019-02/2019-01_aihealth.pdf?subscribe=Download+the+Report)
4. Calo C. J Ethical Implications of Using the Paro Robot with a Focus on Dementia Patient Care, Association for the Advancement of Artificial Intelligence. 2011
5. Choe J, Coffman B.A., Bergstedt D.T Transcranial Direct Current Stimulation Modulates Neuronal Activity and Learning in Pilot Training // Frontiers in Human Neuroscience. February
6. EPI Report // Education inequalities at the school starting gate, September 27, 2017 URL: <https://www.epi.org/publication/education-inequalities-at-the-school-starting-gate>
7. EU guidelines on ethics in artificial intelligence: context and implementation // EPRS, September 2019 URL: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS\\_BRI\(2019\)640163\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf)
8. Frauenberger K. Developing technology for autism: an interdisciplinary approach // Personal and Ubiquitous Computing, 2011, 16(2), pp 117127
9. Furey H, Martin F AI Education Matters: A Modular Approach to AI Ethics Education // AI Matters, 2018, Volume 4, Issue 4, December, pp 13-15
10. Goel, A. K, Joyner, D. A Using AI to teach AI: lessons from an online AI class // AI Mag. 2017, 38, pp 4858
11. Hargreaves, D. Educational research and evidence-based educational practice: A response to critics // Research Intelligence, 1996, 58, pp 1216
12. Huang W.C, Hwang, J.Y Review of affective computing in education/learning: Trends and challenges // British Journal of Educational Technology, 2016, 47(6), pp 1304-1323
13. Laar E., Deursen J, Dijk J, Haan J The relation between 21st-century skills and digital skills: A systematic literature review // Computers in Human Behavior, 2017, Volume 72, July, pp 557-588, Elsevier
14. Laura R.S, Chapman A The technologisation of education: philosophical reflections on being too plugged // Int J Child Spiritual. 2009, 14(3), pp 289298
15. MIT Lab Report An Ethics of Artificial Intelligence Curriculum for Middle School Students // Blakeley H.P., MIT Media Lab, August 2019 URL: <https://www.media.mit.edu/projects/ai-ethics-for-middle-school/overview/>
16. Soller A. A machine learning approach to assessing knowledge sharing during collaborative learning activities // International Society for Learning Sciences, 2012, January
17. Susskind R., Susskind D. The Future of the Professions // Oxford University Press, 2016
18. The Institute for Ethical Ai & Machine Learning // URL: <https://ethical.institute>

19. Yonhap news agency // URL: <https://en.yna.co.kr/view/AEN20191127004800315>



# Этические ориентиры интеллектуальных технологий в образовании

**Горман А. В.**

*МГУ им. М.В. Ломоносова  
Российская Федерация, Москва*

## **Аннотация**

Если мы ставим благополучие каждого отдельного учащегося превыше всего, очевидно, что применение искусственного интеллекта (ИИ) способно серьезно улучшить ситуацию с качеством и равенством доступа к образованию в долгосрочной перспективе. Однако это возможно тогда и только тогда, когда в процессе внедрения и применения таких систем, права и свободы каждого человека не будут нарушены, что требует соответствующего воспитания и образования граждан, консенсуса об этических нормах, их внедрения на регулятивном межправительственном, технологическом и экономических уровнях, а также направленных превентивных технологических решений.

**Ключевые слова:** искусственный интеллект, этика технологий, образование, образовательные технологии

**Дата публикации:** 19.12.2019

## **Ссылка для цитирования:**

Горман А. В. Этические ориентиры интеллектуальных технологий в образовании // Искусственные общества. 2019. Т. 14. Выпуск 4 [Электронный ресурс]. Доступ для зарегистрированных пользователей. URL: <https://artsoc.jes.su/s207751800007918-8-1/> (дата обращения: 28.09.2020). DOI: 10.18254/S207751800007918-8